

BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 09 OCT. 2003

Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

Martine PLANCHE

DOCUMENT DE PRIORITÉ

PRÉSENTÉ OU TRANSMIS
CONFORMÉMENT À LA
RÈGLE 17.1.a) OU b)

INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE

SIEGE
26 bis, rue de Saint Petersburg
75800 PARIS cedex 08
Téléphone : 33 (0)1 53 04 53 04
Télécopie : 33 (0)1 53 04 45 23
www.inpi.fr

26bis, rue de Saint-Petersbourg
75800 Paris Cédex 08
Téléphone: 01 53.04.53.04 Télécopie: 01.42.94.86.54

Code de la propriété intellectuelle-livre VI

REQUÊTE EN DÉLIVRANCE

DATE DE REMISE DES PIÈCES: 27 sept. 2002 N° D'ENREGISTREMENT NATIONAL: 0211952 DÉPARTEMENT DE DÉPÔT: 75 DATE DE DÉPÔT: 27 SEP. 2002	Michel BRAUN THOMSON multimedia 46 Quai Alphonse Le Gallo 92648 BOULOGNE cedex France
Vos références pour ce dossier: PF020128	

1 NATURE DE LA DEMANDE

Demande de brevet

2 TITRE DE L'INVENTION

PROCÉDE DE REGROUPEMENT D'IMAGES D'UNE SEQUENCE VIDEO

3 DECLARATION DE PRIORITE OU REQUETE DU BENEFICE DE LA DATE DE DEPOT D'UNE DEMANDE ANTERIEURE FRANCAISE

Pays ou organisation Date N°

4-1 DEMANDEUR

Nom	THOMSON LICENSING S.A.
Suivi par	BRAUN Michel
Rue	46 Quai Alphonse Le Gallo
Code postal et ville	92100 BOULOGNE-BILLANCOURT
Pays	France
Nationalité	France
Forme juridique	Société anonyme
N° SIREN	383 461 191
Code APE-NAF	322A
N° de téléphone	01 41 86 50 00
N° de télécopie	01 41 86 56 34
Courrier électronique	braunm@thmulti.com

5A MANDATAIRE

Nom	BRAUN
Prénom	Michel
Qualité	Liste spéciale, Pouvoir général: 9016
Cabinet ou Société	THOMSON multimedia
Rue	46 Quai Alphonse Le Gallo
Code postal et ville	92648 BOULOGNE cedex
N° de téléphone	01 41 86 52 68
N° de télécopie	01 41 86 56 34
Courrier électronique	braunm@thmulti.com

6 DOCUMENTS ET FICHIERS JOINTS		Fichier électronique	Pages	Détails
Description		desc.pdf	10	
Revendications		✓	3	6
Dessins		✓	2	4 fig., 1 ex.
Abrégé		✓	1	
Figure d'abrégé		✓	1	fig. 3; 2 ex.
Désignation d'inventeurs				
Listage des sequences, PDF				
Rapport de recherche				
7 MODE DE PAIEMENT				
Mode de paiement	Prélèvement du compte courant			
Numéro du compte client	626			
Remboursement à effectuer sur le compte n°	626			
8 RAPPORT DE RECHERCHE				
Etablissement immédiat				
9 REDEVANCES JOINTES	Devise	Taux	Quantité	Montant à payer
062 Dépôt	EURO	35.00	1.00	35.00
063 Rapport de recherche (R.R.)	EURO	320.00	1.00	320.00
Total à acquitter	EURO			355.00
10 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE				
Signé par	Michel BRAUN			

La loi n°78-17 du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire. Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

L'invention concerne un procédé et un dispositif de regroupement
5 d'images clés par utilisation d'attributs spatiaux-temporels.

Le regroupement, plus connu sous l'appellation anglaise de
« clustering », a pour but de grouper des données en exploitant des mesures
de distances ou similarités entre celles-ci, des données faiblement distantes
étant regroupées au sein d'une même classe. Une application relative aux
10 données vidéo numériques concerne la construction automatique de
résumés vidéo.

La figure 1 représente un schéma général d'un procédé de
construction d'un résumé vidéo d'une séquence d'images.

15 Dans une première étape 1, la séquence vidéo est découpée en
plans vidéo. Une deuxième étape 2 extrait des images caractéristiques ou
images clés des différents plans de la séquence. Pour chaque image clé,
c'est l'étape 3, une signature est calculée, par exemple à partir d'attributs de
l'image tels que la couleur, la structure.... L'étape suivante 4 a pour but de
20 regrouper les différents plans qui se ressemblent en groupes de plans ou
« clusters ». Une mesure de similarité est ainsi effectuée, à partir d'un calcul
quelconque de distances entre les signatures des images clés associées à
chaque plan. L'étape suivante 5 fabrique le résumé en extrayant une image
clé par groupe de plans.

25 Ce procédé permet de réduire le nombre de plans caractéristiques,
par exemple lors de séquences vidéo avec plans alternés, pour créer des
résumés vidéo, des index ...

Une approche connue s'attaquant aux problèmes de regroupement
30 de plans pour la construction de résumés vidéo est celle proposée par Yeung
et Yeo dans le document "Segmentation of video by clustering and graph
analysis", Computer Vision and Image Understanding, vol 71, no 1, July, pp
94-109, 1998. A chaque couple de plans est associée une distance qui est
une mesure sur les différences entre signatures. Outre la distance entre
35 signatures, la méthode proposée s'attache à ne pas rassembler de plans
dont la distance temporelle est supérieure à un seuil temporel T. L'idée sous-

jacente repose sur l'hypothèse que des plans appartenant à une même unité sémantique ne peuvent être très distants. Cette hypothèse présente également l'avantage de limiter le nombre de regroupements potentiels et ainsi de limiter le coût calculatoire. L'algorithme de regroupement utilisé
 5 fonctionne sur des regroupements successifs de plans en commençant par les plus semblables jusqu'à ce que les distances soient toutes supérieures à un seuil.

Dans cette méthode, si deux plans sont séparés de plus de T images, le regroupement n'est pas possible. Dans l'article cité, T est fixé à
 10 une valeur de l'ordre de quelques milliers d'images. Le problème principal de cette méthode réside dans le fait que ce seuil est fixe, qu'il a une incidence importante sur le résultat final et qu'il est de ce fait difficile à fixer a priori. Par exemple, si une scène de dialogue dure plus de 3000 images et si cette valeur dépasse le seuil temporel, un surdécoupage a lieu. Tous les plans
 15 ainsi regroupés doivent être proches deux à deux à la fois visuellement et temporellement. La taille des groupes ainsi générés est de ce fait relativement limitée. Le résultat final de cet algorithme de clustering se caractérise par l'obtention de groupes relativement uniformes en nombre de plans, ce nombre étant généralement faible et de toutes façons limité par la
 20 valeur du seuil T .

L'invention a pour but de pallier les inconvénients précités. Elle a pour objet un procédé de regroupement d'images d'une séquence vidéo constituée de plans et représentée par une structure en graphe, un nœud du
 25 graphe représentant un plan ou une classe de plans définie par des images clés et les nœuds étant reliés par des arêtes, caractérisé en ce qu'il comporte l'itération suivante :

- sélection d'une arête a_k reliant les nœuds n_i et n_j
- calcul du potentiel du nœud n_m , fusion des deux nœuds n_i et n_j , en
 30 fonction des distances entre les attributs des images clés définissant la classe de plans du nœud n_i et ceux des images clés définissant la classe de plans du nœud n_j et en fonction de la distance temporelle de ces images clés,
- calcul du potentiel de chaque arête reliant le nœud fusionné n_m à un
 35 autre nœud du graphe précédemment relié au nœud n_i ou n_j , en fonction des distances entre les attributs des images clés définissant la classe de

plans du nœud fusionné et ceux des images clés définissant la classe de plans de l'autre nœud et en fonction de la distance temporelle entre ces images clés, la nouvelle classe de plans associée au nœud fusionné comprenant les images clés des classes de plans des nœuds fusionnés,

- 5 - fusion des deux nœuds et validation du nouveau graphe si l'énergie de ce graphe, qui est la somme des potentiels des nœuds et des arêtes, est inférieure à l'énergie du graphe avant la fusion.

10 Selon une mise en œuvre particulière de l'invention, le graphe est initialisé en affectant un nœud à chaque plan et en créant des arêtes d'un nœud à un autre nœud si les plans relatifs à ces nœuds sont séparés d'un nombre maximal prédéterminé T de plans.

15 Selon une mise en œuvre particulière de l'invention, la valeur de potentiel associé aux arêtes et aux nœuds est une fonction respectivement décroissante et croissante de la distance temporelle.

20 Le calcul de similarité des images prend en compte la distance entre signatures, par exemple la différence entre histogrammes, et la distance temporelle entre images clés définissant les plans. Cette contrainte temporelle est adaptative, deux plans trop éloignés initialement pour être regroupés pouvant se retrouver potentiellement « regroupables » en cas de fusion de plans intermédiaires.

25 L'algorithme de calcul est basé sur une minimisation énergétique par fusion s'appuyant sur une structure de graphe. Contrairement à la méthode précédente, pour appartenir à un groupe d'images, aussi appelé par la suite classe d'images ou classe de plans, une image candidate doit être proche visuellement et temporellement d'au moins une image du groupe.

30 Un avantage de l'invention est la prise en compte des images même éloignées temporellement, du fait de la fusion, tout en favorisant le regroupement d'images temporellement proches.

35 Un autre avantage réside dans le fait que seule une image d'un groupe doit être similaire à l'image candidate pour autoriser la fusion. Les groupes ainsi créés peuvent comporter plus d'éléments que dans le cas de la méthode classique. Le schéma proposé permet de générer des groupes plus hétérogènes, moins dépendants de valeurs de seuils donc plus proches des données, sans pour autant rendre les calculs trop complexes ou trop longs.

D'autres particularités et avantages de l'invention apparaîtront clairement dans la description suivante donnée à titre d'exemple non limitatif et faite en regard des figures annexées qui représentent :

- la figure 1, un schéma général de la construction de résumés vidéo,
- 5 - la figure 2a, un exemple de graphe initial,
- la figure 2b, le résultat de ce graphe après une itération,
- la figure 3 un organigramme de minimisation de l'énergie du graphe.

10 La méthode proposée s'appuie sur une structure de graphe. Un graphe est constitué d'un ensemble de nœuds et d'arêtes. Un nœud n_i symbolise un plan ou groupe de plans vidéo similaires suivant la distance visuelle et temporelle choisie, chaque plan P_i étant représenté par au moins une image clé. Les arêtes relient les nœuds entre eux, deux nœuds non
15 reliés par une arête n'étant pas potentiellement fusionnables.

Une taille de fenêtre temporelle T est choisie. Elle est donnée en nombre maximal de plans T séparant deux plans dans la séquence vidéo. Lors de l'initialisation du graphe, des arêtes sont créées à partir d'un nœud donné correspondant à un plan donné vers chacun des nœuds dont les
20 plans se trouvent dans la fenêtre temporelle de ce plan donné.

D'une manière pratique, les nœuds sont d'abord reliés par des arêtes selon le déroulement séquentiel de la séquence vidéo. Ce cheminement est ensuite pris en compte, lors de l'adjonction des arêtes complémentaires, pour l'application de la fenêtre temporelle. Alors, la valeur
25 entière T choisie pour cette étape d'initialisation du graphe signifie que seuls les nœuds séparés au maximum de T nœuds du nœud considéré sont pris en compte pour le calcul des arêtes initiales partant de ce nœud. Une arête initiale a_{ij} n'est créée entre le nœud n_i et le nœud n_j que si la distance temporelle d_T séparant ces nœuds auxquels sont associés les plans P_i et P_j
30 est inférieure à un seuil T . Cette contrainte n'est effective qu'à l'initialisation du graphe, les fusions ultérieures de nœuds permettant de relier des nœuds beaucoup plus éloignés à l'origine, si nécessaire.

Le choix de ce seuil à l'initialisation est fait dans un but d'optimisation du temps de calcul très dépendant du nombre d'arêtes.

35 L'initialisation du graphe est effectuée en créant un nœud par plan vidéo. La figure 2a représente un graphe initial constitué d'un ensemble de 6

plans vidéo symbolisés par les nœuds 1 à 6. Dans l'exemple de la figure 2a, à titre d'illustration simplifiée, le seuil temporel ou fenêtre temporelle à l'initialisation est fixé à $T=1$. Le seuil est en réalité choisi égal à 6 dans notre implémentation. Dans l'exemple donc, les arêtes sont créées entre un nœud et les nœuds séparés de ce nœud d'un plan au maximum. La figure 2b, qui correspond à la fusion des nœuds 2 et 4 pour donner le nœud 7 montre cependant que, dès la première itération, les nœuds 6 et 7, reliés par des arêtes, sont en dehors de la fenêtre temporelle.

L'énergie du système est égale à la somme des potentiels des arêtes et des nœuds. Le but est de regrouper ou fusionner les nœuds de manière à diminuer l'énergie globale du système.

Pour définir l'énergie globale du système, la notion de potentiel V_n de nœud et V_a d'arête est introduite. Un potentiel est calculé à partir d'une distance de base entre deux images I_i et I_j , distance temporelle et distance entre attributs.

La valeur du potentiel associé à une arête a_{ij} est définie à l'aide d'une fonction décroissante de la distance. Plus les distances temporelles et les distances entre attributs sont importantes, plus le potentiel est faible. L'objectif global de l'algorithme étant de minimiser l'énergie globale du système, plus les plans attachés à deux nœuds seront proches selon les distances précédemment définies, plus les potentiels des arêtes seront élevés et plus la fusion et donc l'élimination des arêtes sera favorisée.

La valeur du potentiel associé au nœud n_k est définie à partir d'une fonction croissante de la distance. Plus les distances temporelles ou les distances entre attributs calculées entre les images appartenant à deux nœuds distincts sont faibles, plus le potentiel associé au nœud fusion est faible, donc plus le système sera enclin à fusionner les deux nœuds pour créer un nœud de faible potentiel.

Soient $V_a(n_i, n_j)$ et $V_n(n_m)$ respectivement les potentiels d'une arête reliant les nœuds n_i et n_j et d'un nœud n_m .

Nous décrivons dans un premier temps le potentiel lié aux arêtes. On calcule ici les potentiels « inter-images » de l'arête pour chacun des couples formés d'une image parmi les images associées à un nœud extrémité de l'arête et d'une image parmi les images associées à un nœud autre extrémité de l'arête. Il s'agit, en d'autres termes, des potentiels calculés à partir des

distances croisées entre les images associées au premier nœud et celles associées au deuxième nœud aux extrémités de l'arête.

Le potentiel $V_a(n_i, n_j)$ d'une arête s'écrit alors en fonction des distances croisées entre les images I_k appartenant à n_i et les images I_l appartenant à n_j . La valeur minimale suivante est calculée :

$$V_a(n_i, n_j) = \min_{I_k \in P_i, I_l \in P_j} (w_{att} \cdot F_a(K_{att}, \tilde{d}_{att}, d_{att}(I_k, I_l)) + w_T \cdot F_a(K_T, \tilde{d}_T, d_T(I_k, I_l)))$$

où :

- 10 P_m représente l'ensemble des images associé au nœud n_m
- F_a est une fonction décroissante
- w_{att} et w_T représentent les poids liés aux attributs et au temps (dans notre implémentation, $w_{att}=3$ et $w_T=1$)
- $d_{att}(I_k, I_l)$ est la fonction distance associée aux attributs
- 15 caractéristiques des deux images
- \tilde{d}_{att} représente la moyenne des distances entre images calculées sur le graphe initial
- $d_T(I_k, I_l)$ est la distance temporelle séparant les deux images
- \tilde{d}_T est égal à T , T représentant le seuil temporel
- 20 K_{att} et K_T sont deux constantes

L'ensemble des images associé au nœud n_m est également appelé classe de plans associée au nœud n_m . Cette classe regroupe l'ensemble des images des classes de plans fusionnés.

- 25 La distance entre attributs peut être prise comme une différence classe à classe entre histogrammes caractérisant chaque image. La distance temporelle est quant à elle définie comme le nombre de plans séparant les deux plans auxquels appartiennent les images I_k et I_l .

Dans l'implémentation, la fonction décroissante F_a est de la forme

30 suivante :

$$F_a(K, \tilde{d}, d) = \frac{1}{1 + e^{\frac{K}{\tilde{d}}(d - \tilde{d})}}$$

Des valeurs de potentiel sont également calculées pour les nœuds du graphe. Lorsque deux nœuds n_i et n_j correspondant aux images-clé I_i et I_j appartenant respectivement aux classes de plan P_i et P_j sont fusionnés pour créer un nouveau nœud n_k , les images I_i et I_j sont intégrées dans la nouvelle classe de plans P_k associée au nœud n_k . Le nombre d'images associées augmente ainsi au fur et à mesure des fusions. D'une manière générale, le nombre d'images associées à un nouveau nœud correspond à la somme des images associées aux nœuds fusionnés générant le nouveau nœud.

Le potentiel d'un nœud fusion d'un premier et deuxième nœud est obtenu à partir des potentiels calculés sur l'ensemble des couples d'images formés d'une image associée au premier nœud et d'une image associée au second nœud. Cette valeur est fonction de la distance entre attributs d_{att} et de la distance temporelle d_T . Il y a autant de potentiels calculés qu'il y a de couples d'images. La valeur choisie est celle qui donne le potentiel minimum.

Le potentiel $V_n(n_m)$ d'un nœud issu de la fusion de deux nœuds n_i et n_j existants est donc défini par la fonction suivante :

$$V_n(n_m = n_i \cup n_j) = \min_{I_k \in P_i, I_l \in P_j} (w_{att} \cdot F_n(K_{att}, \tilde{d}_{att}, d_{att}(I_k, I_l)) + w_T \cdot F_n(K_T, \tilde{d}_T, d_T(I_k, I_l)))$$

Le potentiel d'un nœud fusion est donc calculé à partir des distances croisées entre les images appartenant à chacun des nœuds candidats à la fusion.

Dans l'implémentation, la fonction croissante F_n est de la forme suivante :

$$F_n(K, \tilde{d}, d) = 1 - \frac{1}{1 + e^{\frac{K}{\tilde{d}}(d - \tilde{d})}} = 1 - F_a(K, \tilde{d}, d)$$

25

A l'initialisation du graphe, le potentiel des nœuds est nul. Il y a une image clé par nœud. C'est au cours du processus de fusion que ce potentiel prend une valeur strictement positive.

Le facteur K utilisé dans les fonctions F_a et F_n est un paramètre qui influence la courbure de la fonction. Plus K est important, plus la fonction tend vers 0 ou vers 1 rapidement quand on s'éloigne de la valeur moyenne \tilde{d} . En terme de potentiel, plus K est important, plus les variations de d autour de la valeur moyenne entraînent des variations importantes du potentiel.

30

Dans le cas de la fonction attachée à la distance entre attributs, K_{att} peut être pris égal à 2. Dans le cas de la fonction attachée à la distance temporelle, K_T peut être pris égal à 7.

5 Lors du calcul d'un potentiel relatif à un couple d'images (I_i, I_j) , plus les images traitées sont éloignées, temporellement ou selon leurs attributs, plus le potentiel de l'arête correspondant à ces deux images est faible, plus le potentiel du nœud correspondant au couple d'images fusionné est fort.

10 La contrainte temporelle apparaît donc sous deux formes dans notre algorithme :

- par l'existence où non d'une arête entre deux nœuds, cette existence étant initialement liée à la distance temporelle séparant deux plans
 - par l'influence de la distance temporelle séparant deux images sur la valeur du potentiel associé aux arêtes ou aux nœuds
- 15

L'énergie globale du système est égale à la somme des potentiels associés à chaque arête et à chaque nœud.

20 Un processus itératif de fusions est mis en œuvre pour obtenir une configuration du système ayant l'énergie globale la plus faible possible. Le but de chaque itération est de faire décroître l'énergie globale. Le calcul de l'énergie minimale nécessiterait d'essayer toutes les combinaisons possibles. Ce nombre pouvant être considérable, en fonction du nombre de plans, le schéma de minimisation énergétique choisi est un schéma sous-optimal. Il ne garantit pas que la solution trouvée corresponde à la configuration minimisant l'énergie totale du système. Il présente cependant l'intérêt d'être rapide tout en fournissant une solution énergétique acceptable.

25

30 Plus la distance entre deux images est faible, plus le potentiel de l'arête relatif à ces deux images est fort donc plus le potentiel de l'arête reliant les nœuds correspondants est élevé. On cherche à diminuer l'énergie globale du système donc à éliminer les arêtes de plus fort potentiel, ce qui revient à éliminer le potentiel relatif à ces arêtes, les potentiels des arêtes sortant des nœuds fusionnés étant recalculés. Ce qui correspond bien à vouloir regrouper les images les plus proches, aussi bien du point de vue attributs que du point de vue temporel.

35

Le potentiel d'un nœud obtenu après fusion des deux nœuds aux extrémités de l'arête supprimée est recalculé. Plus les distances entre les images associées aux nœuds à fusionner sont grandes, plus les potentiels du nœud fusionné relatifs aux couples d'images sont importants, plus le potentiel de ce nœud qui est la valeur minimum parmi ces potentiels est important.

Dans l'exemple de la figure 2b, les nœuds 2 et 4 sont fusionnés pour donner le nœud 7. Les images clé des nœuds 2 et 4 sont donc maintenant associées au nœud 7. Le potentiel du nœud 7 est calculée ainsi que les potentiels des nouvelles arêtes partant de ce nœud, définissant ainsi la nouvelle configuration.

Le processus itératif de réduction d'énergie mis en œuvre est décrit à la figure 3.

La première étape référencée 6 vérifie si toutes les arêtes du graphe ont été testées.

Dans l'affirmative, les itérations sont terminées et les plans ou images clés se trouvent regroupées en classes de plans, créées au fur et à mesure des fusions des nœuds du graphe.

Dans la négative, l'étape suivante est l'étape 7 qui prend en compte une arête a_k non testée du graphe dont les extrémités sont les nœuds que l'on appellera n_i et n_j . Les arêtes sont par exemple sélectionnées en fonction de leur valeur énergétique dans l'ordre décroissant des valeurs du potentiel associé. En effet une arête de fort potentiel supprimée entraîne a priori une baisse énergétique importante.

L'étape suivante 8 effectue un calcul de l'énergie du graphe en supposant une fusion des nœuds n_i et n_j . En plus du potentiel du nœud fusionné, le potentiel des nouvelles arêtes reliant le nouveau nœud à l'ensemble des nœuds adjacents aux nœuds n_i et n_j est pris en compte pour le calcul de l'énergie du nouveau graphe. Si l'énergie trouvée est supérieure à l'énergie avant fusion, test effectué à l'étape 9, on retourne à l'étape 6 puis à l'étape 7 pour traiter l'arête suivante, dans le cas où toutes les arêtes n'ont pas été testées.

Dans le cas où l'énergie calculée est inférieure à l'énergie du graphe avant fusion, l'étape suivante est l'étape 10 qui valide la fusion des nœuds n_i et n_j et le nouveau graphe qui en découle.

5 L'étape suivante 6 vérifie si toutes les arêtes ont été testées. Il s'agit des arêtes non modifiées par la fusion des nœuds mais également des nouvelles arêtes calculées lors de cette fusion. Dans l'affirmative, le graphe final est celui validé à l'étape précédente 10.

10 Une simplification consiste à utiliser un critère d'arrêt sous optimal permettant d'éviter le parcours exhaustif des arêtes. Cette simplification consiste en l'arrêt de la minimisation dès que la fusion potentielle de deux nœuds entraîne une augmentation de l'énergie. Cette simplification est rendue acceptable par la sélection intelligente des arêtes dans l'ordre décroissant de leur valeur de potentiel.

15 Le schéma présenté est globalement sous-optimal du fait du caractère non réversible de la fusion. Le passage à une minimisation de l'énergie optimale peut être envisagé moyennant un coût calculatoire élevé.

20 Le choix de la fonction d'énergie et des potentiels que nous venons de définir nous permet de simplifier le schéma de minimisation énergétique. En effet, le fait de choisir le minimum des potentiels fait que si le potentiel associé à un nœud provenant de la fusion de deux nœuds est inférieur à la somme des potentiels de ces deux nœuds alors la fusion des deux nœuds entraîne nécessairement une baisse du potentiel des arêtes voisines. On peut donc se contenter d'une étude locale de la variation d'énergie, tout en
25 assurant une baisse globale de cette énergie.

Le procédé décrit permet d'obtenir un processus de fusion performant malgré la limitation calculatoire. Celle-ci concerne les possibilités d'appariements initiaux limitées, le nombre d'arêtes étant limité à
30 l'initialisation. Elle concerne également l'exploitation d'une fenêtre réduite à l'initialisation. Cependant, du fait du processus de fusion, des plans éloignés peuvent être regroupés. Les potentiels prennent en compte les distances temporelles et entre attributs. Ils autorisent des fusions si les images clés, bien que distantes, sont très ressemblantes ou bien, si les images clés bien
35 que moyennement ressemblantes, sont très proches temporellement.

REVENDECATIONS

- 5 1 Procédé de regroupement d'images (4) d'une séquence vidéo constituée de plans et représentée par une structure en graphe, un nœud du graphe représentant un plan ou une classe de plans définie par des images clés et les nœuds étant reliés par des arêtes, caractérisé en ce qu'il comporte l'itération suivante :
- 10 - sélection d'une arête a_k (7) reliant les nœuds n_i et n_j
 - calcul du potentiel du nœud n_m (8), fusion des deux nœuds n_i et n_j , en fonction des distances entre les attributs des images clés définissant la classe de plans du nœud n_i et ceux des images clés définissant la classe de plans du nœud n_j et en fonction de la distance temporelle de ces images clés,
- 15 - calcul du potentiel de chaque arête (8) reliant le nœud fusionné à un autre nœud du graphe précédemment relié aux nœuds n_i ou n_j , en fonction des distances entre les attributs des images clés définissant la classe de plans du nœud fusionné et ceux des images clés définissant la classe de plans de l'autre nœud et en fonction de la distance temporelle entre ces images clés, la nouvelle classe de plans associée au nœud fusionné comprenant les images clés des classes de plans des nœuds fusionnés,
- 20 - fusion des deux nœuds et validation du nouveau graphe (10) si l'énergie de ce graphe, qui est la somme des potentiels des nœuds et des arêtes, est inférieure à l'énergie du graphe avant la fusion (9).
- 25
- 30 2 Procédé selon la revendication 1, caractérisé en ce que le graphe est initialisé en affectant un nœud à chaque plan et en ce que des arêtes sont créées d'un nœud à un autre nœud si les plans relatifs à ces nœuds sont séparés d'un nombre maximal prédéterminé T de plans.

3 Procédé selon la revendication 1, caractérisé en ce que la valeur de potentiel associé aux arêtes et aux nœuds est une fonction respectivement décroissante et croissante de la distance temporelle.

5 4 Procédé selon la revendication 2, caractérisé en ce que le potentiel d'une arête reliant deux nœuds n_i et n_j est calculé selon l'équation suivante :

$$V_a(n_i, n_j) = \min_{I_k \in P_i, I_l \in P_j} (w_{att} \cdot F_a(K_{att}, \tilde{d}_{att}, d_{att}(I_k, I_l)) + w_T \cdot F_a(K_T, \tilde{d}_T, d_T(I_k, I_l)))$$

10

et en ce que le potentiel d'un nœud né de la fusion de deux nœuds n_i et n_j est défini par la fonction suivante :

$$V_n(n_m = n_i \cup n_j) = \min_{I_k \in P_i, I_l \in P_j} (w_{att} \cdot F_n(K_{att}, \tilde{d}_{att}, d_{att}(I_k, I_l)) + w_T \cdot F_n(K_T, \tilde{d}_T, d_T(I_k, I_l)))$$

15 où :

P_i représente la classe de plans associée au nœud n_i et comportant les images I_k

F_a est une fonction décroissante

F_n est une fonction croissante

20 w_{att} et w_T représentent les poids liés aux attributs et au temps

$d_{att}(I_k, I_l)$ est la fonction distance associée aux attributs caractéristiques des deux images

\tilde{d}_{att} représente la moyenne des distances entre images calculées sur le graphe initial

25 $d_T(I_k, I_l)$ est la distance temporelle séparant les deux images

\tilde{d}_T est égal à T qui représente le seuil temporel

K_{att} et K_T sont deux constantes

30 5 Procédé selon la revendication 4, caractérisé en ce que la fonction décroissante est de la forme

$$F_a(K, \tilde{d}, d) = \frac{1}{1 + e^{\frac{K}{\tilde{d}}(d - \tilde{d})}}$$

et en ce que la fonction croissante est de la forme

$$F_n(K, \tilde{d}, d) = 1 - \frac{1}{1 + e^{\frac{K}{\tilde{d}}(d - \tilde{d})}}$$

- 5 6 Procédé selon la revendication 1, caractérisé en ce que les itérations sont arrêtées dès que la fusion potentielle de deux nœuds (8) entraîne une augmentation de l'énergie (9), les arêtes étant sélectionnées (6), pour le calcul des fusions, dans l'ordre décroissant de leur valeur de potentiel.

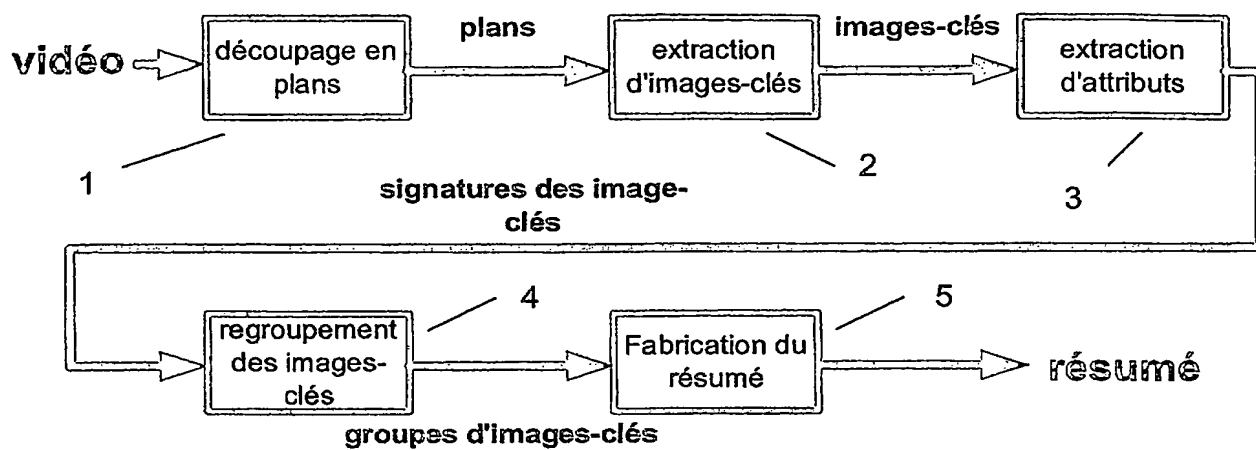


FIG.1

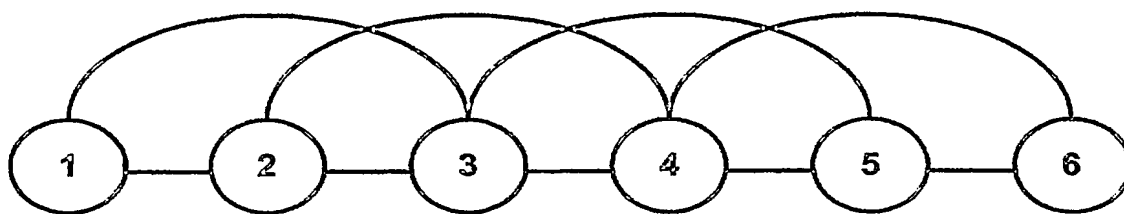


FIG.2a

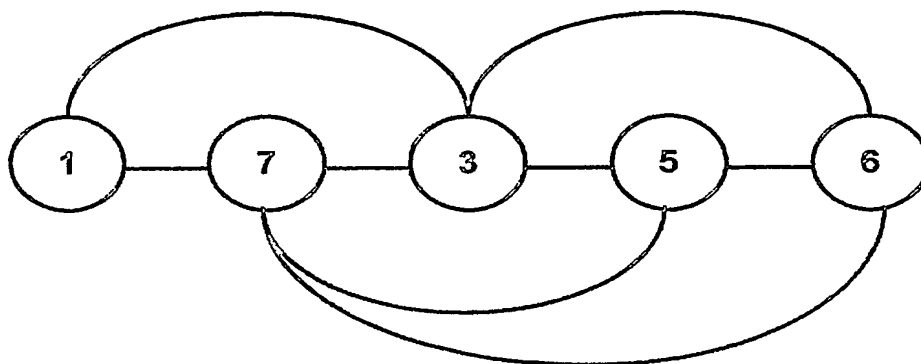


FIG.2b

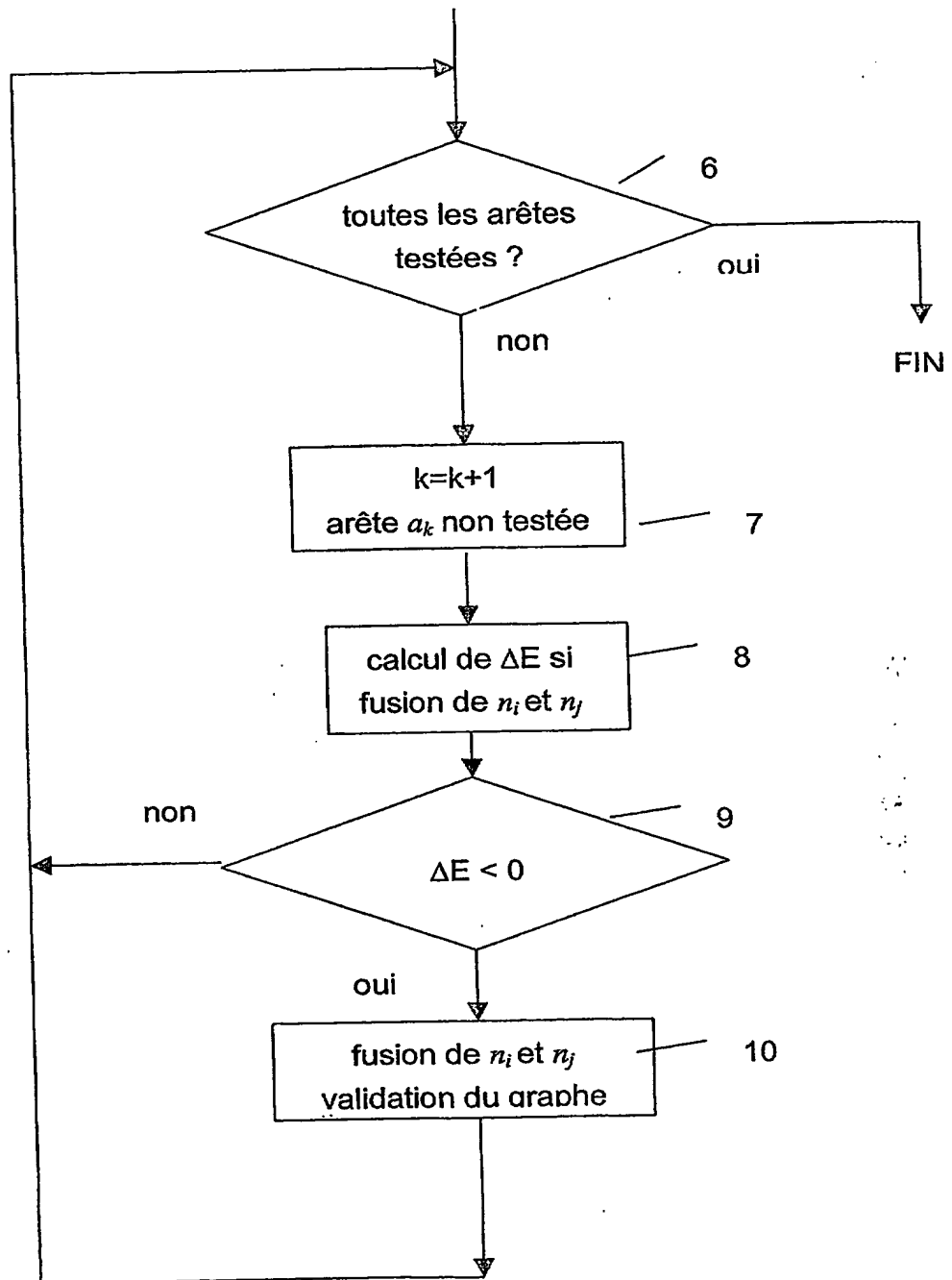


FIG.3

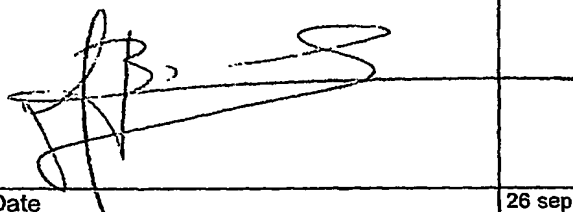


BREVET D'INVENTION

Désignation de l'inventeur

Vos références pour ce dossier	PF020128
N°D'ENREGISTREMENT NATIONAL	0211952
TITRE DE L'INVENTION	
	PROCEDE DE REGROUPEMENT D'IMAGES D'UNE SEQUENCE VIDEO
LE(S) DEMANDEUR(S) OU LE(S) MANDATAIRE(S):	Michel BRAUN

DESIGNE(NT) EN TANT QU'INVENTEUR(S):	
Inventeur 1	
Nom	OISEL
Prénoms	Lionel
Rue	Les Aulnais
Code postal et ville	35137 PLEUMELEUC
Société d'appartenance	THOMSON multimedia
Inventeur 2	
Nom	LE CLERC
Prénoms	François
Rue	21, rue F. Tanguy-Pringent
Code postal et ville	35000 RENNES
Société d'appartenance	THOMSON multimedia
Inventeur 3	
Nom	CHUPEAU
Prénoms	Bertrand
Rue	12, Square de Tanouarn
Code postal et ville	35700 RENNES
Société d'appartenance	THOMSON multimedia

DATE ET SIGNATURE(S) DU (DES) DEMANDEUR(S) OU DU MANDATAIRE	
Signé par:	Michel BRAUN
	
Date	26 sept. 2002

PCT Application
FR0350051

